PERSEREC

Technical Report 15-03
October 2015

Indicators of Suicide Found on Social Networks: Phase 1

Christina M. Hesse
Northrop Grumman Technical Services

Craig Bryan, PsyD, ABPP
National Center for Veterans Studies
University of Utah

Andrée E. Rose
Defense Personnel and Security Research Center
Defense Manpower Data Center

Approved for Public Distribution: Distribution Unlimited
Defense Personnel and Security Research Center
Defense Manpower Data Center

Technical Report 15-03 **October 2015**

Indicators of Suicide Found on Social Networks: Phase 1

Christina M. Hesse—*Northrop Grumman Technical Services*

Craig Bryan, PsyD, ABPP—*National Center for Veterans Studies, University of Utah*

Andrée E. Rose—*Defense Personnel and Security Research Center/DMDC*

Released by – Eric L. Lang

BACKGROUND

Previous social media research found that people publicly post comments about feeling sad, angry, hopeless and suicidal. To date, there have been very few empirical studies examining the prevalence of these types of posts and who is most likely to post this kind of information.

This retrospective study randomly selected 700 military Service personnel who died by suicide and a demographically matched control group (n=700) made up of military Service members who died by causes other than suicide in 2010 and 2011. The data were drawn from the Veteran's Affairs/Department of Defense Suicide Data Repository, a mortality database. Social media checks were conducted on these subjects and all publicly available social networking posts made within a year prior to subjects' deaths were coded for clinical indicators of suicide and analyzed to determine if there were statistically significant relationships between cause of death and the indicators.

HIGHLIGHTS

This report is the first in a two-part series exploring the potential to incorporate social media data into suicide prevention efforts. This report found that this type of data can provide insight into subjects' cognitions and behaviors in the months before their deaths.

Looking at each subject's publicly available online data posted within a year of their deaths, seven indicators differentiated the suicide and control groups. Subjects with posts coded for hopelessness, social withdrawal, and insomnia were more likely to be in the suicide group. Limiting the scope of the analysis to 30 days prior to death revealed that subjects with posts coded for religious affiliation, interpersonal/relationship problems and general distress were more likely to be in the suicide group.

Recommendations for integrating social media data into suicide prevention efforts are provided, as well as recommendations for additional research using social media data.

Defense Personnel and Security Research Center
Defense Manpower Data Center
400 Gigling Rd.
Seaside, CA 93955

REPORT DOCUMENTATION PAGE

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE: 07-Oct-2015	2. REPORT TYPE Technical Report 15-03	3. DATES COVERED: Oct 2014-Oct 2015

4. Indicators of Suicide Found on Social Networks: Phase 1	5a. CONTRACT NUMBER:
	5b. GRANT NUMBER:
	5c. PROGRAM ELEMENT NUMBER:
6. AUTHOR(S): Christina M. Hesse, Craig Bryan, Andrée E. Rose	5d. PROJECT NUMBER:
	5e. TASK NUMBER:
	5f. WORK UNIT NUMBER:
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Defense Personnel and Security Research Center Defense Manpower Data Center 400 Gigling Road Seaside, CA 93955	8. PERFORMING ORGANIZATION REPORT NUMBER PERSEREC: Technical Report 15-03
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSORING/MONITOR'S ACRONYM(S)
	11. SPONSORING/MONITOR'S REPORT NUMBER(S):

12. DISTRIBUTION/AVAILABILITY STATEMENT: (A) Distribution Unlimited

13. SUPPLEMENTARY NOTES:

ABSTRACT: This is a retrospective study examining social media posts made by military Service personnel 1 year prior to their deaths. The sample consisted of 700 randomly selected military Service personnel who died by suicide and a control group that demographically matched the suicide group. Only 482 subjects were found to have social networking profiles, and of those, 315 subjects had publicly available information available for assessment. The social media data were coded by staff and graduate students at the University of Utah's National Center for Veteran's Studies (NCVS) for clinical indicators of suicide and then analyzed to identify relationships between indicators and cause of death. Findings suggest that social media data provide insight on subjects' thoughts and behaviors within the year preceding death.

14. SUBJECT TERMS:

15. SECURITY CLASSIFICATION OF: Unclassified			16. LIMITATION OF ABSTRACT:	17. NUMBER OF PAGES: 59	19a. NAME OF RESPONSIBLE PERSON: Eric L. Lang, Director
a. REPORT: Unclassified	b. ABSTRACT: Unclassified	c. THIS PAGE: Unclassified			19b. TELEPHONE NUMBER (Include area code): 831-583-2846

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI td. Z39.18

PREFACE

In 2013, the Defense Personnel and Security Research Center (PERSEREC) began studying how publicly available social media data could inform suicide intervention and prevention strategies. This effort examines publicly available social media data associated with military Service personnel who died by suicide in 2010 and 2011, and military Service personnel who died by accidental or health related causes during the same time frame. The data are coded for clinical indicators of suicide and analyzed to determine the relationship between cause of death and these indicators.

In a follow-up to this study, we will present findings from a qualitative analysis that identified risk and protective factors, and other contextual factors that were not coded for in this study.

Using findings from these two phases of research, recommendations for the logical next steps for continued research and integration of social media data into intervention and prevention approaches will be offered.

Eric L. Lang, Ph.D.
Director, PERSEREC

EXECUTIVE SUMMARY

INTRODUCTION

Suicide is the second leading cause of death among Americans between the ages of 25 and 34 and is the 10th leading cause of death for all Americans (CDC, 2013 and afsp.org, 2014). In 2013, the rate of suicide in the United States was 12.6 deaths per 100,000 people (afsp.org). With respect to military personnel, the 2013 suicide rate was 18.7 per 100,000 for active duty personnel, 23.4 per 100,000 for Reservists, and 28.9 per 100,000 for members of the National Guard (Department of Defense Suicide Event Report, 2013).

In Fiscal Year 2014, the Defense Suicide Prevention Office (DSPO) funded the Defense Personnel and Security Research Center (PERSEREC), a division within the Defense Manpower Data Center (DMDC), to conduct research on social networking posts made by military Service personnel who died by suicide. PERSEREC collaborated with the National Center for Veteran's Studies (NCVS) to examine whether military Service personnel provide suicide risk indicators on their social networking profiles, and if these indicators can be used to predict suicide. After the posts were coded and initial analyses conducted, additional research questions were formed to include:

(1) What type of person posts publicly available social media data?

(2) What are the differences in posts made immediately prior to suicide? versus over a longer period of time?

(3) What is the impact of 3rd party posts on the predictive model?

METHODOLOGY

This is a retrospective study that drew a random sample of 700 military Service personnel who died by suicide between January 1, 2010 and December 31, 2011, and a demographically matched control group of 700 military Service personnel who died by means other than suicide during the same timeframe. The sample was drawn from 2010 and 2011 Suicide Data Repository (SDR).

A social media data provider used the personal identifiers (e.g., name, date of birth, address, and date of death) maintained in the SDR to conduct Internet searches. Subjects' publicly available social networking posts were collected and de-identified. The NCVS coded these posts for 36 clinical indicators of suicide. The codes were analyzed to determine if the indicators could differentiate between the two groups.

FINDINGS

Of the 1,400 subjects in the sample, 482 subjects had detectable social networking profiles. Of these, 315 subjects had publicly available social media data.

Subjects Coded for at Least One of the 36 Clinical Indicators of Suicide

Subjects with social networking posts coded for one or more clinical indicators were significantly younger than subjects who were not coded for any clinical indicators. Furthermore, subjects with at least one indicator were more likely to be junior enlisted and less likely to have dependents. Alcohol use was the most common indicator, accounting for 24% of all the coded posts.

Differentiating Between the Suicide and Non-suicide Group

Two demographic characteristics were significantly related to cause of death: (1) Religious affiliation and (2) Marital status. Subjects who died by suicide were less likely to self?-identify as Christians and were more likely to be married.

Independently, none of the clinical indicators were related to cause of death. However, five of the indicators were suggestive of significant relationships with cause of death. Odds ratios indicate a larger sample size would have likely produced a significant relationship between cause of death and the following indicators: suicidal rehearsal and preparatory behaviors, medication misuse, feelings of being trapped or helpless, withdrawing from others, and perceived burdensomeness.

Analyzing all Social Media Data Posted within the Year before Death

All social networking posts made by subjects and third parties (i.e., online friends) within the 1-year period before subjects' deaths were analyzed and revealed religious affiliation (demographic variable), hopelessness, social withdrawal, insomnia, sarcasm, physical problems, and anxiety optimally differentiated subjects in the suicide group from subjects in the non-suicide group.

Analyzing all Social Media Data Posted within the Month before Death

A second analysis examined social media data generated within the 30-day window before death to determine if any indicators may serve as flags for near-term or "imminent" suicide risk. The results indicate that during the month immediately preceding their deaths, users whose social networking profiles included more content characterized by interpersonal/relationship problems and general distress were more likely to be suicide cases whereas users whose social networking profiles included more content characterized by thwarted belongingness and anger were less likely to be suicide cases.

Analyzing Subject-generated Data Posted within the Year before Death

A third analysis removed third-party posts from the dataset and focused exclusively on subject-generated posts made within the 1-year window before death. The purpose of this analysis was to determine if third-party posts affected the relationship between the clinical indicators and cause of death. Religious affiliation,

a demographic variable, and wish for death, a clinical indicator, were the only variables that differentiated the two groups.

Analyzing Subject-generated Data Posted within the Month before Death

During the month immediately preceding their deaths, subjects who posted more content characterized by a significant loss and alcohol use were more likely to be in the suicide group, whereas users who posted more content characterized by anger were less likely to be in the suicide group.

DISCUSSION

The primary purpose of this study was to evaluate social media data belonging to military Service personnel who died by suicide to find out if they were posting content containing indicator(s) of suicide prior to death. The results from this effort demonstrate that social media data are a valuable source of information because it offers a unique experience of seeing and reading Service members' behaviors, activities, and thoughts preceding their death.

When examining posts made by subjects and third parties, findings suggest that in the year prior to their deaths, Service members who died by suicide are more likely to have a more pessimistic outlook in life and/or to be surrounded by a social network that communicates a more pessimistic worldview. Those who died by suicide were also more likely to avoid interpersonal situations and/or lack of interest in participating in activities with others, and had more frequent conversations about sleep problems.

Refining the scope to 30 days prior to death, results suggest that in the period of time immediately preceding a Service member's death by suicide, they tended to communicate more often about difficulties related to interpersonal relationships and tended to express generalized stress with greater frequency. In contrast, Service members who died by suicide were less likely to communicate feelings of anger, which may suggest the Service member had "resigned" themselves to their situation.

When removing third-party data from the regression analysis, regardless of time, new indicators emerge, indicating that third-party data affects the relationship between cause of death and the clinical indicators. This finding suggests that any research involving social media data should carefully consider how third-party data are integrated into the research.

RECOMMENDATIONS

Recommendations for Suicide Prevention and Intervention

(1) This study should be replicated using a larger sample size for the purpose of detecting additional statistically significant relationships between cause of

death and clinical indicators. Additional subjects could be drawn from the 2012, 2013, and 2014 cohorts with the Suicide Data Repository (SDR).

(2) DSPO should explore options for integrating publicly available social media data into its Wellness Assessment Risk Nexus.

(3) Because social networking posts can sometimes provide details about subjects' lives in the days and months before their deaths, DSPO and the military Service's suicide prevention offices should consider testing social media data as a new data source for psychological autopsies.

Recommendations for Social Media Research

(4) Current processes for flagging concerning social media data involve both automation and manual review. To create greater time and resource efficiencies, additional research on natural language processing and sentiment analysis is necessary to enhance automated detection and response to online indicators of intent to die by suicide.

(5) Recognizing that social media user behaviors evolve and new social media platforms emerge over short periods of time, it will be important to continue studying this field of communication for user cues related to overall mental health wellness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF TABLES IN APPENDICES

INTRODUCTION

NOTE: Refer to APPENDIX A: Term Definitions for definitions of key terms used throughout this report.

In Fiscal Year 2014, the Defense Suicide Prevention Office (DSPO) funded the Defense Personnel and Security Research Center (PERSEREC), a division within the Defense Manpower Data Center (DMDC), to conduct research on publicly available social networking posts made by military Service personnel who died by suicide. PERSEREC collaborated with the National Center for Veteran's Studies (NCVS) to examine whether military Service personnel provide suicide risk indicators on their social networking profiles, and if these indicators can predict suicide. After the social media data were collected, coded, and analyzed, additional research questions were formed to include:

(1) What type of person posts publicly available social media data?

(2) What are the differences in posts made immediately prior to suicide versus over a longer period of time?

(3) What is the impact of 3rd party posts on the predictive model?

The purpose of this report is to present the findings from this effort and provide recommendations for integrating suicide prevention and intervention into social media.

BACKGROUND

The DoD has devoted resources to reduce the number of suicides within the military. DoDD 6490.14, *Defense Suicide Prevention Program,* states the DoD must take substantial efforts to reduce suicide rates among Service members, and the Services must identify members at risk for suicide and evaluate the efficacy of suicide prevention programs. In 2011, the department established DSPO to serve as the oversight authority for suicide prevention efforts in the military. DSPO's vision is to "enable total force fitness through suicide prevention and resilience programs and policies to ensure Service members and their families overcome risk factors and are mission ready from entry on duty to retirement or separation" (DSPO, n.d.). One facet of suicide prevention is early detection of risk factors.

Clinical Indicators of Suicide

Indicators of suicide include individual, relational, and societal concerns. These factors may not be direct causes of suicidal thoughts and behaviors; however, they are associated with death by suicide. Suicidal indicators are either long-standing vulnerabilities or acute risk factors present during an active suicidal episode. The fluid vulnerability theory, which guided the clinical coding and analyses component of this study, categorizes indicators of suicide according to this dichotomy (Rudd, 2006).

According to the theory, some individuals have long-standing vulnerabilities (e.g., history of psychiatric illness, pessimistic outlook, and trauma history) that put them more at risk for suicide. The theory defines acute risk factors as psychiatric and behavioral signs and symptoms present during an active suicidal episode (e.g., agitation, sleep disturbance, and social withdrawal). The two types of risk factors are organized into four domains representing various ways in which a suicidal episode manifests: (1) behavioral, (2) physical, (3) emotional, and (4) cognitive. In addition, trigger events and external stressors also affect an individual's risk for suicide. Long-standing vulnerabilities, acute risk factors, triggering events, and external stressors all play a critical role in understanding and assessing an individual's risk to die by suicide.

Suicide Prevention and Social Media

The 2012 National Strategy for Suicide Prevention recognized the importance of online communication by including it as an objective, declaring, "Increase communication efforts conducted online that promote positive messages and support safe crisis intervention strategies" (U.S. Surgeon General and the National Action Alliance for Suicide Prevention, 2012). However, there are few empirical studies examining online expressions of suicidal ideation or pre-cursors to ideation.

A project funded by the Office of the Director of National Intelligence (ODNI) analyzed social media data to determine its utility as a data source for security clearance determinations. Researchers found that social media data provided insight into subjects' mental health and that this data source might be useful for suicide prevention research. The project collected and analyzed publicly available online content associated with 3,370 cleared Army personnel. At least five subjects had social media posts indicating a desire to engage in self-harm, intent to die by suicide, or other severe expressions of depression (Rose & Whiteley, 2014). For example, one subject wrote a nonfiction book detailing personal experiences with statutory rape and attempted suicide, while another subject posted statements expressing suicide ideation, anger, and despair. Even in cases without explicit indicators of suicide ideations, social media posts and open source data revealed information related to a number of suicide risk factors including financial issues, drug involvement, psychological conditions, and criminal conduct.

Recent studies investigated if social media data could be useful in suicide prevention and detection. A study by Cash et al. (2013) found adolescents, ages 13-24, use the social networking website MySpace[1] to share comments about their

---

[1] One of the main limitations of this study is that it focused exclusively on MySpace and excluded some of the more heavily used social networks and microblogs (e.g., Facebook and Twitter). New social media platforms are frequently released and users' behaviors evolve over time (Rose, Hesse, & Garcia, 2014). These conditions will limit the relevance of all social media research and encourage continued study within this field.

suicidal thoughts. Some of the discovered topics included: relationships, mental health, substance use and abuse, and different methods of suicide.

Another study examined Twitter conversations for suicide-related posts to determine if these posts were related to suicide rates. Researchers identified search terms related to suicide risk factors and warning signs and used this to identify at-risk suicide Twitter conversation (tweets deemed to be jokes, sarcasm, or nonpertinent were removed) (Jashinksy, Burton, Hanson, West, Giruad-Carrier, Barnes, & Argyle, 2013). Some of the tweet categories included depressive feelings (e.g., *I feel so worthless today*), psychological disorders (e.g., *…what to say but yes, I have been diagnosed with anorexia since late 2009 and early 2010*), and family violence or discord (e.g., *BIGGEST fight with dad EVER. Ended in a fist fight, I've packed my bags & I'm leaving. I hold a grudge so dunno how long b4 we talk again*) (Jashinksy, Burton, et. al, 2013). Using the geolocations of the posts, tweets were grouped by state and compared to actual suicide rates. Analyses revealed an association between rates of tweets related to suicide and actual suicide statistics. States with the highest rates of suicide (i.e., Midwestern and Western states and Alaska) reported high proportions of suicide-related tweeters (Jashinksy et al., 2013).

Social media platforms have also been examined for signs of suicide risk using predictive linguistic-driven models. For the Durkheim project, researchers used clinical notes from the U.S. Department of Veteran's Affairs (VA) to develop linguistics-driven prediction models to identify words and combinations of words associated with suicide (Poulin et al., 2014). Models with word pairs had better predictive accuracy than models with single words (an average score of 64% compared to 59%). The goal of this research would be to apply these algorithms to social media data and ultimately develop an effective suicide intervention strategy.

For now, online social media prevention efforts are focused on making information available to the individual in need and providing online mechanisms for users to report their "friends'" suicidal indicators. Facebook recently collaborated with other Internet companies to develop a list of best practices to address the spill-over of suicide into cyberspace (Curry, 2012). Additionally, Whisper, a mobile application, allows users to anonymously post "secrets" accompanied by an image, recently launched its nonprofit website Your Voice as a digital platform for users to openly discuss suicide and other mental health issues. The website was created in response to users posting about self-harm and suicide on Whisper (Buhr, 2014). The goal is for Your Voice to serve as a platform for individuals who are struggling with various mental health issues to find information and resources, as well as read testimonials of people with similar experiences. Social media is a growing resource for suicide intervention and prevention strategies.

Online Communities and Suicide Prevention

One important aspect of social networking sites (SNS) is the interpersonal networks they create and their ability to bring individuals together. As a result of this characteristic, social networks are sometimes used as a way to reach out to suicidal individuals for intervention. The organization "Battle in Distress" is a crises response team that uses Facebook and other social media platforms to connect U.S. veterans and Service members "who are in or at risk for being in mental, emotional, financial, or psychological distress to the individuals, organizations, government entities, or other entities that provide services or assistance that can ameliorate the underlying problems that lead them to a state of distress" (Battle in Distress, n.d.). The organization strives to change how society views and responds to mental health issues affecting Service members and veterans. One of the main services the group provides is the Crisis Response Team which monitors SNSs 24/7 and uses social media to help veterans and Service members in need.

The group was created in response to a message posted on Facebook, and the responses generated (Battle in Distress, n.d.). Daniel Caddy, a National Guard Staff Sergeant, used Facebook to seek help for a Soldier who expressed intent to die by suicide. Caddy posted a plea on a community Facebook page for Service members asking for help to locate and intervene. Caddy asked if any members were in the Kingsport, Tennessee area, or had a battle[2] in the area who could reach out to the soldier. He went on to explain he was having difficulty contacting the soldier's chain of command and there was a serious concern the soldier was going to "take his own life" (Basu, 2013). Within minutes of posting this message, people started responding and sharing the post. Using geolocation from the soldier's phone, he was located and his unit commanders found him alone and intoxicated in his room. After seeing how effective social media was as an intervention tool, Caddy created the "Battle in Distress" Facebook page and website.

While social media was extremely successful in this incident, there are cases where Service personnel have expressed their suicidal thoughts and ideations online, but intervention was not successful. Prior to his death, Pvt. Jordan DuBois posted "I'm goin [sic] to kill myself this is my last post…miss u [sic] all…" as well as a photograph with the caption "My last picture" (Watts, 2012). Family, friends, and other soldiers tried to reach DuBois and locate him, but unfortunately they were unable to reach him in time. His death was determined to be a suicide.

In another case, Daniel Rey Wolfe, a Marine veteran, also used Facebook to express his suicidal ideations. Wolfe announced his plans to die by suicide on Facebook, and then posted images documenting his death. He posted four images on the day of his death including an image of a half-empty bottle of vodka and Jack Daniel's on the floor with a handwritten note that said "ROT IN." He captioned this photo

[2] Battle is a term often used by Service members to describe a fellow service member.

"Byee bitches." The other three photographs were images of a leg and arm with numerous cuts and scratches with the caption "Is it real yet fuckers" (Weinstein, 2014). In addition to posting the images of his suicide, he also posted status updates with indicators of his desire to die by suicide. Two of these updates read: "The only fight I ever lost was the one to myself" and "When my body moves no more give me a Vikings funeral" (Weinstein, 2014). Friends commented on the pictures in an attempt to reach out and support Wolfe, and tried to use Facebook to locate him. Battles in Distress contacted local police and hospitals trying to find him, but unfortunately they were all too late.

It is evident that social media may provide a unique insight into an individual's thoughts and behaviors prior to their death. This effort explores the usefulness of social media data for identifying clinical indictors of intent to die by suicide.

METHODOLOGY

The following section describes the study's sample, data collection methods, and methods used for analyzing the coded social media data. PERSEREC selected the study's sample and sent identifiable information to the social media data provider. The data provider collected subjects' publicly available social media data and sent de-identified reports to researchers at the NCVS, located at the University of Utah. NCVS researchers coded the data for indicators of suicide and performed statistical analysis on the data. PERSEREC conducted descriptive analyses on the data, and performed the qualitative content analysis.

SAMPLE

The subjects included in this effort were drawn from the Suicide Data Repository (SDR), a repository containing data from the National Death Index (NDI) and the Defense Casualty Analysis System (DCAS). The total sample used in this effort consists of 1,400 military Service personnel, 700 who died by suicide and 700 who died from reasons other than suicide (see APPENDIX B for the list of causes of death). Individuals from all Service Components, as well as the Reserves, National Guard, and Coast Guard, were included. Both groups were demographically matched on age, gender, marital status, number of dependents, and rank.

The groups consisted mostly of men (non-suicide 92.0%, n=644; suicide 94.1%, n=659), ranging in age from 17 to 80 years old (non-suicide M=30.46, SD=10.95; suicide M=29.64, SD=9.32). Nearly half of the subjects were identified as white (non-suicide 43.0%, n=301; suicide 45.7%, n=320) but almost half of the subjects did not have a listed race[3] (non-suicide 43.3%, n=303; suicide 46.9%, n=328). A little more than two-thirds of the subjects' highest level of education was high school (non-suicide 70.9%, n=496; suicide 72.7%, n=509) and nearly half of the subjects did not have dependents (non-suicide 46.9%, n=328; suicide 47%, n=329). See Table 1 for a more extensive list of demographic details.

[3] All demographic data were collected from DMDC's data files. Researchers did not have control over missing demographic data.

Table 1
Descriptive Statistics

Sample Characteristics	Non-suicide n	Non-suicide % of 700	Suicide n	Suicide % of 700
Gender				
• Male	644	92.0	659	94.1
• Female	56	8.0	41	5.9
Marital Status				
• Never married	331	47.3	322	46.0
• Married	303	43.3	333	47.6
• Divorced	52	7.42	37	5.3
• Legally separated	2	0.3	5	0.7
• Widowed	1	0.1	1	0.1
• Unknown	11	1.6	2	0.3
Dependents				
• None	328	46.9	329	47.0
• 1	133	19.0	129	18.4
• 2	92	13.1	93	13.3
• 3	83	11.9	73	10.4
• 4 or more	49	7.0	63	9.0
• Unknown	15	2.1	13	1.9
Race				
• Caucasian	301	43.0	320	45.7
• Black or African American	74	10.6	36	5.1
• American Indian/Alaskan Native	13	1.9	10	1.4
• Asian	7	1.0	6	0.9
• Other	2	0.3	0	0.0
• Unknown	303	43.3	328	46.9
Education				
• High school (HS)	496	70.9	509	72.7
• College	105	15.0	79	11.3
• Some college, no HS diploma	50	7.1	59	8.4
• Less than HS	21	3.0	18	2.6
• Graduate school	14	2.0	17	1.4
• Unknown	14	2.0	18	2.6
Religious Affiliation				
• Christian	426	60.9	360	51.4
• No religion	165	23.6	216	30.9
• Other	19	2.7	29	4.1
• Unknown	90	12.9	95	13.9

Approximately half of the subjects were Junior Enlisted Service members (non-suicide 50.6%, *n*=354; suicide 53.1%, *n*=372), and 36.1% (*n*=253) of the non-suicide group and 36.4% (*n*=255) of the suicide group were Non-commissioned Officers

(NCO). A little over half of the subjects were in the Army (non-suicide 59.7%, *n*=418; suicide 61.1%, *n*=428), and 13.6% (*n*=95) of the non-suicide group and 16.0% (*n*=112) were in the Air Force. There were 282 (40.3%) regular Service members in the non-suicide group and 319 (45.6%) regular Service members in the suicide group. See Table 2 for a more in depth analysis of subjects' rank, Service, and Component.

Table 2
Military Status

	Non-suicide		Suicide	
	n	% of 700	n	% of 700
Military Component				
Regular	282	40.3	319	45.6
Reserves	234	33.4	221	31.6
Guard	184	26.3	160	22.9
Military Branch				
Army	418	59.7	428	61.1
Air Force	95	13.6	112	16.0
Navy	84	12.0	82	11.7
Marine Corps	88	12.6	67	9.6
Coast Guard	14	2.0	11	1.6
Public Health	1	<1.0	0	-
Service Rank				
Junior Enlisted	354	50.6	372	53.1
NCO	253	36.1	255	36.4
Officer	67	9.6	49	7.0
Senior Enlisted	15	2.1	19	2.7
Warrant Officer	11	1.6	5	<1.0

SOCIAL MEDIA DATA COLLECTION

The social media data provider utilized subjects' identifiable information to conduct automated searches for subjects' publicly available online content. For this effort, publicly available refers to information that is not meaningfully restricted. The social media data were not password protected and did not require log-in or special access. Ideally, the data provider would have had access to the subjects' email addresses because it is a unique online identifier often linked to social networking profiles, but the SDR does not maintain this information. Therefore, the social media searches had to be conducted using only personal, real-world identifiers.

To collect the data, the provider utilized automated methods to search and aggregate online data. During this process, a proprietary identity resolution algorithm was applied to ensure all data were related to the subject. This process returned a series of links, but only data meeting the following criteria were collected:

(1) Social media data came from one of the following types of websites:

- Social Networking (e.g., Facebook, Myspace, LinkedIn)
- Microblogging (e.g., Twitter, Pinterest, Reddit)
- Blogging (e.g., Tumblr, LiveJournal)

(2) Social media data were only collected if it was posted between the date of death and 1 year prior to the date of death.[4] However, there were many instances in the dataset where, after death, third parties posted comments on pictures or status updates subjects made before their deaths.

A social media report was generated on each subject in the sample, and all social media data that were within scope were included in the report. The report only contained de-identified data. Names and faces were redacted by the data provider.

CODING SOCIAL MEDIA DATA FOR INDICATORS OF SUICIDE

Researchers at the NCVS coded all social media posts to determine the presence or absence of 36 potential indicators of suicide (see APPENDIX C for a list of indicators and the corresponding definitions). The 36 indicators were organized into five clusters, consistent with the fluid vulnerability theory of suicide (Rudd, 2006): triggers, cognitions or thoughts, behaviors, physical/physiological symptoms, or emotions. The clusters were defined as followed:

(1) *Triggers* entail descriptions of stressful situations or life circumstances that could potentially activate an acute suicidal crisis.

(2) *Cognitions or thoughts* entail descriptions or verbalizations of beliefs, assumptions, and subjective appraisals of the self and/or a situation.

(3) *Behaviors* entail descriptions of observable actions.

(4) *Physical/physiological symptoms* entail descriptions of somatic complaints or health-related issues.

(5) *Emotions* entail descriptions of feelings or internal affective states or experiences.

[4]A 1-year time range was selected because researchers wanted to focus specifically in the months leading up to subjects' deaths.

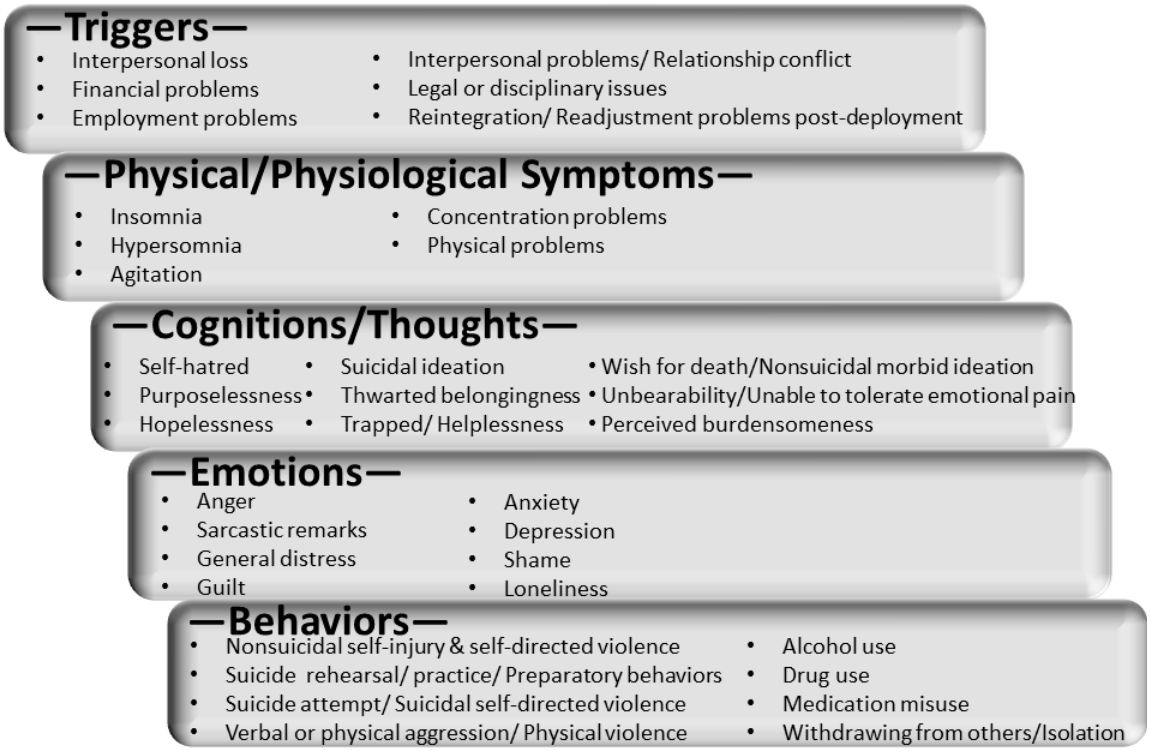

Figure 1 NCVS Coding Clusters and Categories

The lead NCVS researcher and project coordinator trained 15 coders to identify each of the 36 codes. This was a labor-intensive process, whereby two research assistants were responsible for reviewing and counting the total number of posts per subject to ensure agreement in total content that was to be coded. Additionally, each subject's social media report was coded by two people. Interrater reliabilities were high; 34 of the 36 indicators achieved kappa coefficients greater than .90. Kappas ranged from .67 to 1.00, with a mean of .97 (SD = .06) and a median of .99. Discrepancies in coding were discussed and resolved by individual team members. Where discrepancies could not be resolved, a final decision was made by an assigned master coder. All coders were blind to subjects' cause of death in an attempt to eliminate any coder bias.

Analyses

Differences Between Subjects with Social Networking Profiles and Subjects without Social Networking Profiles

The primary objective of this effort was to determine if publicly available social media data associated with military personnel provide signals of one's trajectory towards dying by suicide. To answer this question, analyses first focused on identifying the types of individuals who use social media and engage with others on social networking sites. Social media research is not necessarily appropriate for all

populations. This analysis identified military Service personnel who are most likely to have at least one social networking profile.

Differences Between Subjects with Coded Social Media Data and Subjects without Coded Social Media Data

Chi-square and Mann-Whitney[5] tests were performed to compare differences between subjects with and without data flagged for one of the 36 indicators of suicide. This provided a better understanding of how subjects with relevant social media data differed from those who did not.

The next set of analyses sought to answer whether there was a difference between the social networking profiles affiliated with subjects in the suicide group and the social networking profile affiliated with subjects in the non-suicide group. Analyses of the indicators were conducted without adjusting for demographics, and then repeated with the following demographic variables entered as covariates: gender, religious affiliation, age, number of dependents, age of enlistment, race, branch of Service, rank, and component.

Stepwise logistic regression with forward selection was used to determine if any of the indicators on social media can differentiate Service members who died from suicide from Service members who died from reasons other than suicide. A significance level of $p<.350$ was specified for the inclusion of a variable in the model and a significance level of $p<.250$ was specified to keep a variable in the model, as recommended by Hosmer and Lemeshow (1989). Stepwise logistic regression is designed to find predictor variables most useful in predicting the outcome. Variables are added to the model one at a time and then tested to determine if its inclusion results in a statistically significant improvement in overall model fit. Variables are dropped from the model if they do not improve model fit whereas variables that improve model fit are retained.

Stepwise logistic regression was selected to identify variables that best differentiated suicides from controls because there was no theoretical or conceptual basis for assuming that certain combinations of predictors would be better predictors of suicide than others.

The next level of analyses focused on identifying other variables that may affect the predictive nature of the data, specifically date of post in relation to date of death and the influence of third-party posts. A stepwise logistic regression was run on only the data posted during the 30 days immediately preceding the subjects' deaths. The intention here was to see if there is any indication of the acute stressors of suicide compared to the long-standing stressors. A stepwise logistic regression was also used to analyze the data with and without third-party posts, in an attempt to examine the influence conversations had on the predictive nature of

[5] The Mann-Whitney test was used because the data were not evenly distributed.

the indicators. Data posted by third-parties provide context for the subject-generated posts, as well as information and/or observations about the subject that would otherwise go unreported. Anything posted on a subject's social networking profile contributes to the overall theme and tone of the profile, even instances where third-parties are posting about their own situation. If researchers were to exclude third-party comments, half of the conversation would be lost and a rich level of detail would be missing from the data.

FINDINGS

DIFFERENCES BETWEEN SUBJECTS WITH AND WITHOUT SOCIAL MEDIA PROFILES

Of the 1,400 subjects in the sample, 482 subjects (34.4%) had a social networking profile (n= 237, 49.2% suicide group, n=245, 50.8% non-suicide group). The data pulled from these profiles included publicly available social media posts.

Statistically significant differences between the subjects who had at least one social networking profile and those who did not have a profile were found for age, if the subject had dependents, marital status, component and rank. Subjects with at least one social networking profile were younger (Median age =25)[6] than subjects who did not have a social networking profile (Median age = 27).[7] Subjects with at least one dependent (including spouses) were significantly less likely to have social media data[8] than those subjects without dependents. Married subjects,[9] subjects in the Regular component,[10] and Junior enlisted[11] were more likely to have a profile, while subjects in the Reserve component were significantly less likely than those in the Regular component or the National Guard to have a profile.

Demographic Characteristics of Subjects with Social Media Data

Of the 482 subjects with at least one social networking profile, 315 subjects had publicly available social media data. The remaining 167 subjects used privacy settings to prevent access to their social media postings or created a social networking profile but never posted anything on it.

CLINCAL INDICATORS

Differences between Subjects with and without Clinical Indicators

Researchers at NCVS reviewed the data and determined that of the 315 subjects with social media data, 107 (34.0%) subjects in the suicide group and 116 (36.8%) subjects in the non-suicide group had at least one of the 36 indicators. Statistically significant differences between the subjects who had at least one indicator of suicide and those who did not were found for age, dependents, and rank. Subjects with at least one indicator of suicide were younger (Median age=23) than the subjects who did not have relevant social media data (Median=25).[12] Subjects who

[6] The data were not evenly distributed and were skewed toward younger subjects. Because of this, the median is a better indicator than the mean.
[7] U=165412.00, z=-7.78, $p<.001$
[8] U=184107.50, z=-4.51, $p<.001$
[9] $x^2(5)= 23.85$, $p<.001$
[10] $x^2(2)=13.56$, $p<.001$
[11] $x^2(4)= 40.00$, $p<.001$
[12] U=23845.50, z=-3.31, $p<.05$

had relevant social media data were also less likely to have dependents[13] than those who did not have coded social media data. Finally, junior enlisted Service members were more likely to have relevant content[14] than other ranks.

Descriptive Analysis of the Clinical Indicators

Suicide attempt appeared least often (*n*=1) and alcohol use was coded for most often (*n*=1,251). See Table 3. The most common indicators, appearing more than once per subject on average, were: alcohol use, general distress, anger, aggression or violence, and physical problems/somatic symptoms.[15]

A few indicators were only present in one group or the other. For example, suicide rehearsal (see Figure 2) was only present in the suicide group (*n*=3); however, suicide attempt (see Figure 3) was only present in the non-suicide group (*n*=1). Although this indicator may be specific to suicide, its considerable infrequency suggests it may not be sufficient as a broad-based signal for suicide. See APPENDIX C for additional descriptive statistics of these indicators.

Subject: Gun loaded safety off so close to pulling the trigger why live when u have nothing to live for right. I love you [name redacted]. I always will

Figure 2 Suicide Rehearsal[16] found on Google+

Subject: I took my leap and found myself hitting the rocks. The sharp pain that goes through my body and bones that are broken. To hear the truth and see that I made a decision turned out to be a bad thing. Now I lay here. Broken, torn, bloodied and full of heartache.

Figure 3 Suicide Attempt[17] Found on Facebook

13 U=25148.00, z=-2.39, p<.05

14 $x^2(4)$= 15.92, *p*<.05

15 The clinical indicators were largely intercorrelated with each other in a positive direction, and when grouped by cluster, the clusters were also intercorrelated All five clusters showed moderate to strong positive correlations with each other. Because the clinical indicators were all established risk factors for suicide, intercorrelations were expected to be moderate to high. To assess for potential multicollinearity, diagnostic tests (i.e., variance inflation factors) were conducted with each analysis. In all cases, variance inflation factors were well below accepted thresholds.

16 Posted on July 21, 2011 and subject's date of death by suicide is July 22, 2011.

17 Posted 4 weeks prior to subject's death in a car accident (occupant).

Table 3
Frequency of Indicators by Non-suicide and Suicide Groups

Indicator (by Domain)	Non-Suicide (n=116) *n*	Non-Suicide (n=116) *% of Total*	Suicide (n=107) *n*	Suicide (n=107) *% of Total*	Total Number of Times an Indicator Appears within the Entire Data Set *n**
Triggers					
• Employment	8	29.6	19	70.4	27
• Financial	14	41.2	20	58.8	34
• Interpersonal/relationship problem	112	36.7	193	63.3	305
• Legal/disciplinary	32	42.1	44	57.9	76
• Reintegration/readjustment	3	50.0	3	50.0	6
• Interpersonal loss	327	67.7	156	32.3	483
Cognitive					
• Feeling trapped or helpless	15	45.2	19	55.9	34
• Hopelessness	10	29.4	24	70.6	34
• Perceived burdensomeness	2	33.3	4	66.6	6
• Purposelessness/ meaninglessness	2	50.0	2	50.0	4
• Self-hatred	5	27.8	13	72.2	18
• Suicide ideation	6	37.5	10	62.5	16
• Thwarted belongingness	6	46.2	7	53.8	13
• Unbearability	14	50.0	14	50.0	28
• Wish for death/nonsuicidal morbid ideation	7	24.1	22	75.9	29
Behavioral					
• Verbal or physical aggression/ Physical violence	221	46.0	259	54.0	480
• Alcohol use	647	53.7	557	46.3	1,204
• Drug use	31	47.0	35	53.0	66
• Medication misuse	4	44.4	5	55.6	9
• Nonsuicidal self-injury	2	50.0	2	50.0	4
• Suicide attempt	1	100.0	0	0.0	1
• Suicide rehearsal, preparation, or practice	0	0.0	3	100.0	3
• Withdrawing from others	2	25.0	6	75.0	8
Physical					
• Agitation	12	66.7	6	33.3	18
• Concentration problems	3	42.9	4	57.1	7
• Hypersomnia / excessive sleep	3	60.0	2	40.0	5

Indicator (by Domain)	Non-Suicide (n=116)		Suicide (n=107)		Total Number of Times an Indicator Appears within the Entire Data Set
	n	*% of Total*	*n*	*% of Total*	*n**
• Insomnia / sleep disturbance	28	37.8	46	62.2	74
• Physical problems /somatic symptoms	256	55.3	207	44.7	463
Emotional					
• Anger	258	42.4	351	57.6	609
• Anxiety	7	53.8	6	46.2	13
• Depression	29	46.0	34	54.0	63
• General distress	304	43.1	402	56.9	706
• Guilt	6	31.6	13	68.4	19
• Loneliness	31	38.3	50	61.7	81
• Sarcasm	188	54.8	155	45.2	343
• Shame	0	0.0	3	100.0	3

*The number of times an indicator appears within a dataset may be more than the total number of subjects, because indicators may have appeared more than once within subjects' social media reports.

Differentiating Between the Suicide and Non-suicide Groups

Demographic Data

Religious affiliation and cause of death were significantly related, suggesting that individuals who died by suicide are somewhat less likely to have identified with a sect of Christianity than either no religious affiliation or one of the other listed religions (e.g. Judaism, Buddhism, Other).[18] Individuals in the suicide group were also more likely to be married.[19]

Indicators of Suicide

Univariate logistic regression[20] analyses revealed that none of the indicators independently differentiated the suicide and non-suicide groups. However, there were several indicators with a large odds ratio,[21] suggesting that if there were higher frequencies of these codes there might have been a significant relationship. These variables were: suicidal rehearsal/preparation/practice (OR=5.54), medication misuse (OR=3.33), feeling trapped or helpless (OR=3.13), withdrawing from others (OR=2.79) and perceived burdensomeness (OR=2.22).

Analyzing all Social Media Data Posted within the Year before Death

A stepwise logistic regression was run to identify the best combination of indicators.[22] First, analyses included all coded data (i.e., subject-generated posts and posts from third parties) to determine if the content contained on users' social networking profiles could differentiate between the suicide and non-suicide groups. The final model, presented in Table 4, was comprised of several variables that optimally differentiated the suicide from the non-suicide group: religious affiliation, hopelessness, social withdrawal, insomnia, sarcasm, physical problems/somatic symptoms, and anxiety. Two of these variables, hopelessness and insomnia, were associated with significantly increased odds for suicide. Specifically, with each additional post containing content about hopelessness, the odds of being in the suicide group increased by a factor of 1.93, and with each additional post

---

[18] $x^2(3)=14.60$, $p<.05$

[19] $x^2(5)=11.58$, $p<.05$

[20] Univariate logistic regression: a regression for a single binary/dichotomous outcome. It is very similar to a linear regression, but the regression weights here refer to changes in the log odds of the outcome, rather than predicting the linear association between a predictor and an outcome (as is the case in linear regression).

[21] Odds ratio: derived from a logistic regression, this tells us the increase or decrease in the odds of an outcome based on a 1-point increase in a predictor. It tells us the strength of the association between a predictor and an outcome, focusing on the changes in odds related to that predictor. These tend to be used to interpret the results of logistic regressions, because the regression weights are not as intuitive.

[22] Indicators that showed large differences between the suicide and non-suicide group were sometimes not significant in the stepwise regression due to small sample sizes for those indicators, or because of a small unique effect of that indicator (that is, some indicators were correlated with each other).

containing content about insomnia, the odds of being in the suicide group increased by 1.60. These results indicate that during the year preceding their deaths, subjects whose social networking profiles included more content characterized by hopelessness and insomnia were more likely to be suicide cases whereas users whose social networking profiles included more content characterized by sarcasm, physical problems/somatic symptoms, social withdrawal, and anxiety were less likely to be in the suicide group.

Table 4
Results of Stepwise Logistic Regression Model Predicting Cause of Death Based on all Available Posts during the 12 Months Preceding Death

					95% Confidence Interval	
	df	**Wald χ^2**	**p**	**OR**	**lower**	**upper**
• Religious affiliation (demographic characteristic)	7	11.72	0.110	--	--	--
• Hopelessness	1	4.49	0.034	1.93	1.05	3.54
• Social withdrawal	1	2.70	0.100	5.17	0.73	36.59
• Insomnia	1	4.98	0.026	1.60	1.06	2.43
• Sarcasm	1	2.85	0.092	0.89	0.77	1.02
• Physical problems/somatic symptoms	1	2.04	0.153	0.93	0.84	1.03
• Anxiety	1	1.83	0.176	0.32	0.06	1.68

Analyzing all Social Media Data Posted within the Month before Death

To determine if any indicators may serve as flags for near-term or "imminent" suicide risk, the data set was restricted to posts that occurred 30 days immediately preceding subjects' deaths. Stepwise logistic regression was repeated. Results are presented in Table 5. One demographic variable and four clinical indicators were included in the final model: religious affiliation, interpersonal/relationship problems, thwarted belongingness, anger, and general distress. Results indicated that the odds of being in the suicide group increased by 1.93 with each additional post about a relationship problem and increased by a factor of 1.52 with each additional post about general distress. In contrast, the odds of being in the suicide group decreased by a factor of 0.31 with each additional post coded for anger.

Table 5
Results of Stepwise Logistic Regression Model Predicting Cause of Death Based on all Available Posts During the 30 Days Preceding Death

					95% Confidence Interval	
	df	Wald x^2	p	OR	lower	upper
• Religious affiliation	7	15.77	0.027	--	--	--
• Interpersonal/relationship problem	1	3.09	0.079	1.93	0.93	4.00
• Thwarted belongingness	1	3.40	0.065	0.09	0.01	1.16
• Anger	1	6.54	0.011	0.31	0.13	0.76
• General distress	1	5.21	0.023	1.52	1.06	2.17

Analyzing Subject-generated Data Posted within the Year before Death

Analyses were run to identify the difference in outcomes when only subject-generated posts were included in the stepwise logistic regression analysis. The purpose of this analysis was to determine if third-party data affected the relationships between cause of death and any of the indicators. The final model, presented in Table 6 was comprised of only two variables: (1) religious affiliation (a demographic characteristic) and (2) wish for death. However, independently, these indicators are not significantly related to the odds of being in the suicide group.

Table 6
Results of Stepwise Logistic Regression Model Predicting Cause of Death Based on Subject-generated posts only During the 12 Months Preceding Death

					95% Confidence Interval	
	df	Wald x^2	p	OR	lower	upper
• Religious affiliation (demographic)	7	10.99	0.139	--	--	--
• Wish for death / nonsuicidal morbid ideation	1	2.49	0.115	1.73	0.88	3.40

Analyzing Subject-generated Data Posted within the Month Before Death

This analysis was repeated with the dataset further restricted to only subject-generated posts that occurred within the 30-day period before death. Results are presented in Table 7. Three indicators were included in the final model: (1) significant loss, (2) alcohol use, and (3) anger. The odds of being in the suicide group increased by a factor of 2.08 with each additional post coded for significant loss. In contrast, the probability of being in the suicide group decreased by a factor of 0.50 with each additional post coded for anger.

Table 7
Results of Stepwise Logistic Regression Model Predicting Cause of Death Based on Subject-generated Data During the 30 Days Preceding Death

	df	Wald x^2	p	OR	95% Confidence Interval lower	95% Confidence Interval upper
• Significant loss	1	5.28	0.022	2.08	1.11	3.89
• Alcohol use	1	3.76	0.053	1.39	1.00	1.93
• Anger	1	4.18	0.041	0.50	0.26	0.97

DISCUSSION

The goal of this study was to evaluate social media data associated with military Service personnel, who died by suicide, and determine if they were posting indicator(s) of suicide prior to their deaths. The results suggest that social media data can be a valuable source of information because it provides the unique experience of seeing and reading Service members' behaviors and thoughts prior to their deaths and, in some instances, provides early signals of one's trajectory towards suicide.

Study Limitations

The social media data provider identified social networking profiles for 30% of the sample. There are three limitations to this study that may account for this finding. First, the most recent data within the SDR are from 2010 and 2011, therefore subjects in the sample have been deceased for at least 4 years. The gap in time between subjects' deaths and data collection may have caused social networking websites to deactivate non-active profiles, or subjects' family members may have deactivated subjects' profiles.

Second, the data provider generally relies on e-mail addresses to help identify social networking profiles. E-mail addresses are a unique online identifier that are often associated with social networking profiles. E-mail addresses are not maintained in the SDR, therefore searches for social networking profiles relied on personal identifiers such as first and last name, gender, age at time of death, military status, place of birth, and place of death. Many more subjects in the sample may have had social networking profiles but they were not detected because the profiles were not linked with enough personal identifiers for the data provider to have a high degree of confidence that they belonged to the subjects of interest.

Lastly, the control group consisted of military Service personnel who died by accidental and health-related causes. A living control group would have been preferred but because of legal (i.e., consent), privacy (limitations on the use of data set forth by the Privacy Act of 1974), and ethical (i.e., the need to intervene if a subject posts threatening content and the complications that could arise from that intervention if it was a false positive) considerations, living subjects were excluded from the sample.

SUBJECTS WITH SOCIAL MEDIA ACCOUNTS

These findings provided some insight into who, among military Service personnel, are more likely to be social media users. Younger, married, junior enlisted Service members in the regular component were more likely to have at least one social networking profile. Identifying the population of social media users could affect how social media data are leveraged in an applied setting.

For example, Nock et al., (2013) reported that current assessment methods were a major challenge to studying suicidal thoughts and behaviors. Before suicide events, it is often difficult for people to describe factors contributing to their behaviors. After suicide events, researchers must rely on psychological autopsies to identify risk factors. Social media data provide an opportunity to enhance psychological autopsies by providing a snapshot of individuals' cognitions and activities in the days, weeks, and months before their deaths. Social media checks are time and resource intensive so performing these checks on all subjects who die by suicide may not be feasible. Greater efficiencies could be achieved by performing these types of checks on subjects who are most likely to be social media users.

CLINICAL INDICATORS OF SUICIDE

The 36 indicators were created based on clinical theories of suicide but the low frequency of these indicators within the social media posts suggests that clinical indicators may not always be appropriate when evaluating social media data. While the way people interact online may have some bearing on how they behave in the physical world, there is a clear separation between the two. People have more control over how they present online, and therefore they might be less inclined to post about suicide attempts or self-harming behavior. However, they may be more likely to post depressing song lyrics as a way to share how they are feeling, or reach out to third parties for interaction, or even post pictures of alcohol.

Alcohol was the most frequently coded indicator and it deserves additional attention. While it is easy to dismiss this indicator because of its prevalence among both the suicide and non-suicide group, alcohol is the second most frequent risk factor for suicidal behavior and its continued abuse can lead to changes in the brain that lead to depression (Center for Disease Control, n.d.). Further research is necessary to explore the nuances of alcohol-related posts because there may be important distinctions that were not addressed in this study. For example, an image of a subject drinking from what appears to be a glass of wine at a wedding may be different than a person's Facebook status update describing one's intent to get "wasted" at a party. The former is a socially acceptable way to share that one is drinking alcohol and may be dismissed, while the latter is suggestive of alcohol abuse. However, without additional contextual information the posts do not provide much detail with respect to subjects' actual use of alcohol.

Differentiating Between the Suicide and Non-suicide Group

Two demographics were significantly related to cause of death: (1) religious affiliation and (2) marital status. Subjects who died by suicide were less likely to identify as Christians and were more likely to be married. Independently, none of the 36 indicators were related to cause of death. However, five of the indicators were suggestive of significant relationships with cause of death. Findings suggest that if the sample was larger analyses may have detected a significant relationship between cause of death and the following indicators: suicidal rehearsal/preparatory

behaviors/practice, medication misuse, feeling trapped or helpless, withdrawing from others, and perceived burdensomeness.

Analyzing all Social Media Data Posted within the Year Before Death

When coded social media data from a year prior to subjects' deaths were analyzed in combinations, a stepwise logistic regression revealed religious affiliation, hopelessness, social withdrawal, insomnia, sarcasm, physical problems, and anxiety differentiated between the subjects in the suicide and non-suicide groups. This may indicate that in the year prior to their deaths, Service members who died by suicide were more likely to have a more pessimistic outlook in life and/or to be surrounded by a social network that communicated a more pessimistic worldview. Those who died by suicide were also more likely to avoid interpersonal situations and/or lack interest in participating in activities with others, and had more frequent conversations about sleep problems.

Analyzing all Social Media Data Posted within the Month Before Death

The results suggest that in the 30 day period of time immediately preceding a Service member's death by suicide, they tended to communicate more often about difficulties related to interpersonal relationships and tended to express generalized stress with greater frequency. In contrast, Service members who died by suicide were less likely to communicate feelings of anger, which may suggest the Service member had "resigned" themselves to their situation and had "given up" their attempts to resolve their life problems or situations. Service members who died by suicide were also less likely to communicate feelings of disconnection or isolation from others, which is notable in light of the finding that posts focused on relationship problems are more common among those who died by suicide. Taken together, these findings might also suggest a form of "resignation" or "giving up," in that Service members no longer feel compelled to talk about their perceived isolation from others despite their interpersonal struggles.

The results found here show evidence of changes in posting behavior over time that might be informative about an individual's level of suicide risk, particularly during periods of time relatively close to their death. In a future study, we plan to examine the nature of posting over the year prior to death to determine if individuals who die by suicide might show a particular pattern of posting behavior over time. This study should inform us whether or not social media might be a useful tool in identifying individuals who are at imminent risk of suicide, or if trajectories of posting behavior over time might help us to identify those individuals likely to be at risk in the future.

Analyzing Subject-generated Data Posted within the Year Before Death

Limiting the analysis to subject-generated posts made within the year before death, religious affiliation and wish for death were the only two indicators, that, when combined, were able to differentiate between the two groups. The differences

between the regression models including and excluding third-party data suggest that this is an important factor to consider when analyzing social media data. While it may appear counterintuitive to include third-party posts about their personal drug use, unemployment, alcohol use, marital discord, etc., online exchanges between the subjects and third parties provide a glimpse of the stressors to which subjects are exposed, even if only by secondary exposure. Nonetheless, research needs to further explore online interactions to understand how users internalize these exchanges and identify other possible effects.

Analyzing Subject-generated Data Posted within the Month Before Death

An analysis including only subject-generated posts 30 days prior to death identified a statistically significant model that included increased probability of being in the suicide group for each additional post coded for significant personal loss and alcohol use. Subjects were less likely to be in the suicide group for each additional post coded for anger.

The differences between the regression models from the 1-year analysis and 30-day analysis, and the analyses including and excluding third-party data, suggest that third-party posts and time are important factors to consider when analyzing social media data. As mentioned earlier, additional research is needed to better understand online social interactions and when these interactions occur with respect to a suicide event.

Conclusion

This research confirmed that some military Service personnel provide clinical indicators of intent to die by suicide on social media. While the results from the current study did not provide strong indicators, they do suggest that additional research, with a larger sample, could identify statistically significant relationships between cause of death and the indicators.

Suicide prevention notwithstanding, military Service personnel in both groups posted comments indicating they were experiencing significant stress and alcohol dependency issues. While some of the subjects with clinical indicators of suicide did not die by suicide, the findings suggest that they were likely struggling with mental health issues. Therefore, it is important to not just think of social media research in terms of suicide prevention, but overall Service member mental health wellness.

RECOMMENDATIONS

Recommendations for Suicide Prevention and Intervention

(1) This study should be replicated using a larger sample size for the purpose of detecting additional statistically significant relationships between cause of death and clinical indicators. Additional subjects could be drawn from the 2012, 2013, and 2014 cohorts with the SDR.

(2) DSPO should explore options for integrating publicly available social media data into its Wellness Assessment Risk Nexus.

(3) Because social networking posts can sometimes provide details about subjects' lives in the days and months before their deaths, DSPO and the military Service's suicide prevention offices should consider testing social media data as a new data source for psychological autopsies.

Recommendations for Social Media Research

(4) Current processes for flagging concerning social media data involve both automation and manual review. To create greater time and resource efficiencies, additional research on natural language processing and sentiment analysis is necessary to enhance automated detection and response to online indicators of intent to die by suicide.

(5) Recognizing that social media user behaviors evolve and new social media platforms emerge over short periods of time, it will be important to continue studying this field of communication for user cues related to overall mental health wellness.

REFERENCES

Agerbo, E., Gunnell, D., Bonde, J. P., Bo Mortensen, P., & Nordentoft, M. (2007). Suicide and occupation: the impact of socio-economic, demographic and psychiatric differences. *Psychological Medicine.* 37(8), pp1131-1140.

Basu, M.B. (2013). Websites become tool for stopping veteran suicides. *CNN*. Retrieved from http://www.cnn.com/2013/09/21/us/facebook-suicide/.

Battle in distress (n.d.) Retrieved from https://www.facebook.com/BattleInDistress/info.

Battle in distress: Success stories. (2013, January 12). Retrieved from http://battleindistress.org/sucess-stories/

Buhr, S. (2014). Whisper's nonprofit your voice is now a digital platform to discuss depression openly. *TechCrunch Network.* Retrieved from http://techcrunch.com/2014/08/14/whisper-switches-its-non-profit-your-voice-to-a-digital-platform-to-give-a-voice-to-those-who-suffer/

Cash, S. J., Thelwall, M., Peck, S. N., Ferrell, J. Z., and Bridge, J. A. (2013). Adolescent suicide statements on MySpace. *Cyberpsychology, Behavior, and Social Networking, 16*(3), pp 166-174.

Center for Disease Control and Prevention (CDC), National Center for Injury Prevention and Control, Division of Violence Prevention. (2012). *Suicide: Facts at a glance.* Retrieved from http://www.cdc.gov/violenceprevention/pdf/suicide_datasheet_2012-a.pdf.

Center for Disease Control. (n.d.). *Suicides due to alcohol and/or drug overdose; a data brief from the National Violent Death Reporting System*. Retrieved from http://www.cdc.gov/ViolencePrevention/pdf/NVDRS_Data_Brief-a.pdf

Chioqueta, A P. & Stiles, T. C. (2007). The relationship between psychological buffers, hopelessness, and suicidal ideation: Identification of protective factors. *The Journal of Crisis Intervention.* 28(2), pp 67-73.

Curry, C. (2012). Facebook taking proactive stance against suicides. *ABC News*. Retrieved from http://abcnews.go.com/US/facebook-social-media-sites-working-offer-suicide-users/story?id=16029673.

Defense Suicide Prevention Office (DSPO). (2014). *Department of Defense quarterly suicide report calendar year 2013: 4th quarter.* Retrieved from http://www.suicideoutreach.org/Docs/Reports/DoD%20Quarterly%20Suicide%20Report%20CY2013%20Q4.pdf.

Defense Suicide Prevention Office (n.d.) Retrieved from http://www.suicideoutreach.org/.

REFERENCES

Department of Defense Directive 6490.14, *Defense Suicide Prevention Program.* June 18, 2013.

Gilman et al. (2014). Sociodemographic and career history predictors of suicide mortality in the United States Army 2004-2009. *Psychological Medicine.*

Hosmer, D. W. & Lemeshow, S. (1989). *Applied logistic regression.* New York: John Wiley & Sons, Inc.

Jashinksy, J., Burton, S.H., Hanson, C.L., West, J., Giraud-Carrier, C., Barnes, M.D., Argyle, T. (2013). Tracking suicide risk factors through twitter in the US. *Crisis.*

LeardMann, C.A. et al. (2013). Risk factors associated with suicide in current and former US military personnel. *JAMA 310*(5).

Mann, JJ, Apter, A, Bertolote, Jl, Beautrais, A., Currier, D., Haass, A., Hergerl, U., Longgvist, J., Malone, K., Marusic, L., Patton, G., Phillips, M., Rutz, W., Rihmer, Z., Schmidtke, A., Shaffer, D., Silverman, M., Takahashi, Y., Varnik, A., Wasserman, D., Yip, P., & Hendin, H. (2005). Suicide prevention strategies: a systematic review. *Journal of the American Medical Association.* 2005; 294: 2064–2074

Nock, M. K., Deming, C. A., Fullerton, C. S., Gilman, S. E., Goldenberg, M., Kessler, R. C., McCarroll, J. E., McLaughlin, K. A., Peterson, C., Schoenbaum, M., Stanley, B., & Ursano, R. J. (2014). Suicide among soldiers: A review of psychosocial risk and protective factors. *Psychiatry.* (76)2: pp 97-125.

Poulin, C., Shiner, B., Thompson, P., Vepstas, L., Young-Xu, Y., Goertzel, B., Watts, B., Flashman, L. & McAllister, T. (2014) Predicting Risk of suicide by analyzing the text of clinical notes. *PLoS ONE* 9(1).

Rose, A.E., & Whiteley, P.J. (2014). *Using social media data to help inform personnel security clearance determinations* (TR 14-02). Monterey, CA: Defense Personnel and Security Research Center/Defense Manpower Data Center. (For Official Use Only).

Rudd, M.D. (2006). *Fluid vulnerability theory: a cognitive approach to understanding the process of acute and chronic risk*. In T. E. Ellis (Ed.), *Cognition and Suicide: Theory, Research, and Therapy* (pp. 355-368). American Psychological Association. Washington, DC:

U.S. Department of Health and Human Services), Office of the Surgeon General and National Action Alliance for Suicide Prevention. (2012). *2012 National Strategy for Suicide Prevention: Goals and Objectives for Action.* Washington, DC: HHS.

U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. (n.d.) VA/DoD clinical practice guideline for assessment and

management of patients at risk for suicide. Retrieved from http://www.guideline.gov/content.aspx?id=47023&search=Risk+factor+

Watts, L. (2012). Ft. Carson soldier posted about suicide before crash: Pvt. Jordan DuBois posted suicide note on Facebook. *KRDO.com.* Retrieved from http://www.krdo.com/news/Ft-Carson-Solider-Posted-About-Suicide-Before-Crash/14816294.

Weinstein, A. (2014). Why Facebook kept up photos of a Marine's bloody suicide. *Gawker.* Retrieved from http://gawker.com/why-facebook-kept-up-photos-of-a-marines-bloody-suicide-1572987701.

APPENDIX A:

TERM DEFINITIONS

- Content: Text, pictures, video, and any other meaningful material that is on the Internet.
- Facebook: A social networking website. Users can create a personal account/profile, add other users as friends, and exchange messages and profile updates.
- "Friend:" Another social media user with whom an individual connects and allows to view their account/profile. Users must request to be someone's friend and then must be accepted by the user.
- Google+: A social networking service operated by Google Inc. The service launched on June 28, 2011 in an invite-only "field testing" phase.
- Microblog: Social media site, such as Twitter, that allows users to share small elements of information such as short sentences, individual images, video and website links.
- MySpace: A social networking website with an emphasis on music.
- Post: Content published on a social networking page or account/profile. May include text, video, images, or links to other web pages.
- Privacy Settings: Options offered by each social media platform to allow users control who can and cannot see their account/profile.
- Profiles: Information that users provide about themselves when signing up for a social networking site as well as a picture and basic information. This may include personal and business interests, a "blurb" and tags to help people search for like-minded people.
- Social Media: Tools and platforms people use to publish, converse, and share content online.
- Social Media Data: Online content that is created and/or curated by users. May include text, videos, images, or links to web pages.
- Social media report: A document that contains all publicly available, de-identified social media data collected by the data provider.
- Social Networks: An online service, platform or site that focuses on facilitating the building of relationships between people who might share interests, activities, backgrounds, or real-life connections.
- Status Update: a feature allowing users to share a short, text-based message with online friends.
- Subject-generated Posts: Social media content created by the subjects in this study's sample.
- Twitter: A platform that allows users to share 140-character-long messages publicly. Users can "follow" each other as a way of subscribing to one another's messages. Additionally, users can use the @username command to direct a message towards another Twitter user.
- Tumblr: A microblogging platform that allows users to post text, photos, videos, links, quotes, and audio to their tumblelog, a short-form blog.

APPENDIX B:

CAUSES OF DEATH

Table B-1
Non-suicide Deaths

Category and Description	**N**
Accident	
Accidental inhalation and ingestion of food or other objects causing obstruction of respiratory tract	4
Air and space transport accidents	56
Cataclysmic storm and flood	5
Fall from one level to another	21
Fall on same level	7
Unspecified fall	8
Motorcyclist involved in any accident except collision with railway train	N
Occupant of car, pickup truck or van involved in collision with other motor vehicle	74
Occupant of heavy transport vehicle or bus involved in collision with other motor vehicle	1
Occupant of motor vehicle involved in collision with other (non-motorized) road vehicle, streetcar, animal or pedestrian	46
Occupant of motor vehicle involved in noncollision accident	55
Occupant of special-use motor vehicle involved in any accident	19
Water transport accidents	8
Assault	
Assault (homicide) by all other and unspecified means and their sequelae	18
Assault (homicide) by bodily force	3
Assault (homicide) by discharge of firearms	112
Assault (homicide) by hanging, strangulation and suffocation	2
Assault (homicide) by sharp object	13
Medical	
Asthma	8
Congestive heart failure	9
Other and unspecified heart failure	6
Obstetric causes	1
Obstetric death of unspecified cause	1
Other deaths related to pregnancy, childbirth and the puerperium	2

Table B-2
Suicide Deaths

Description	N
Intentional self-harm (suicide) by all other and unspecified means and their sequelae	19
Intentional self-harm (suicide) by discharge of firearms	475
Intentional self-harm (suicide) by hanging, strangulation and suffocation	150
Intentional self-harm (suicide) by jumping from a high place	7
Intentional self-poisoning (suicide) by and exposure to drugs and other biological substances	33
Intentional self-poisoning (suicide) by and exposure to other and unspecified solid or liquid substances and their vapors	3
Intentional self-poisoning (suicide) by and exposure to other gases and vapors	13

APPENDIX C:

DESCRIPTIVE STATISTICS FOR THE 36 INDICATORS OF SUICIDE

Descriptive statistics for each of the 36 indicators are reported in Table C-1.

Table C-1
Descriptive Statistics for the 36 Indicators of Suicide

Category/Indicator (by Domain)	N[a]	M[c]	SD[d]	Max[e]
Triggers				
Employment problems	27	0.13	0.55	29
Financial problems	34	0.15	0.60	34
Interpersonal problems/relationship conflict	305	1.37	3.97	27
Legal/disciplinary issues	76	0.34	1.09	10
Reintegration/readjustment problems	6	0.03	0.19	2
Interpersonal loss	483	2.22	5.02	49
Cognitive				
Trapped/helplessness	35	0.16	0.67	6
Hopelessness	34	0.15	0.66	5
Perceived burdensomeness	6	0.03	0.16	1
Purposelessness /meaninglessness	5	0.02	0.18	2
Self-hatred	18	0.08	0.36	3
Suicide ideation	16	0.07	0.29	2
Thwarted belongingness	13	0.06	0.37	4
Unbearability/Unable to tolerate emotional pain	28	0.13	0.42	3
Wish for death/nonsuicidal morbid ideation	29	0.13	0.50	3
Behavioral				
Verbal of physical aggression/physical violence	490	2.20	4.98	33
Alcohol use	1251	5.61	12.04	91
Drug use	66	0.30	1.04	10
Medication misuse	9	0.04	0.32	4
Non-suicidal self-injury/Non-suicidal self-directed violence	4	0.02	0.13	1
Suicide attempt/suicidal self-directed violence	1	-	0.067	1
Suicidal rehearsal/practice/preparatory behaviors	4	0.02	0.16	2
Withdrawing from others/isolation	8	0.04	0.21	2
Physical				
Agitation	18	0.08	0.41	4
Concentration problems	7	003	0.18	1
Hypersomnia	5	0.02	0.15	1

Category/Indicator (by Domain)	N[a]	M[c]	SD[d]	Max[e]
Insomnia	75	0.34	1.12	14
Physical problems	467	2.09	5.43	45
Emotional				
Anger	620	2.78	5.59	39
Anxiety				
Depression	63	0.28	1.17	14
General distress	723	3.24	8.23	88
Guilt	19	0.09	0.35	3
Loneliness	81	0.36	0.99	7
Sarcasm	346	1.55	3.59	39
Shame	3	0.01	0.12	1

APPENDIX D:

HIGHLIGHTING AND REDACTING

The social media data provider redacted all identifiable information throughout the data collection phase. Personally identifiable information was defined as:

(1) Full name

(2) Mailing or Home Address

(3) Email address

(4) Date of Birth

(5) Telephone number

(6) Username/Internet handles

(7) Nickname

Some information was not redacted: explicit statements of sex (male and female), pronouns (him, her, he, she), military status, educational information, general location information, employment information, and marital status. Independently, these pieces of information were not enough to identify the subjects. Additionally, first names were not redacted unless it appeared in combination with a middle or last name. Images were de-identified by placing a black box over all visible faces.